Assessing the performance of eight real-time updating models and procedures for the Brosna River

M. Goswami, K. M. O'Connor, K. P. Bhattarai, A. Y. Shamseldin

Hydrology and Earth System Sciences, 9(4), *394–411* (2005) © EGU

Assessing the performance of eight real-time updating models and procedures for the Brosna River

M. Goswami[1], K.M. O'Connor[1], K.P. Bhattarai[1] and A.Y. Shamseldin[2]

[1]Department of Engineering Hydrology, National University of Ireland, Galway, Ireland

[2]Department of Civil and Environmental Engineering, The University of Auckland, Private Bag 92019, Auckland, New Zealand

Email for corresponding author: kieran.oconnor@nuigalway.ie

Abstract

The flow forecasting performance of eight updating models, incorporated in the **G**alway River **F**low **M**odelling and **F**orecasting **S**ystem (GFMFS), was assessed using daily data (rainfall, evaporation and discharge) of the Irish Brosna catchment (1207 km^2), considering their one to six days lead-time discharge forecasts. The *Perfect Forecast of Input over the Forecast Lead-time* scenario was adopted, where required, in place of actual rainfall forecasts. The eight updating models were: (i) the standard linear Auto-Regressive (AR) model, applied to the forecast errors (residuals) of a simulation (non-updating) rainfall-runoff model; (ii) the Neural Network Updating (NNU) model, also using such residuals as input; (iii) the Linear Transfer Function (LTF) model, applied to the simulated and the recently observed discharges; (iv) the Non-linear Auto-Regressive eXogenous-Input Model (NARXM), also a neural network-type structure, but having wide options of using recently observed values of one or more of the three data series, together with non-updated simulated outflows, as inputs; (v) the Parametric Simple Linear Model (PSLM), of LTF-type, using recent rainfall and observed discharge data; (vi) the Parametric Linear perturbation Model (PLPM), also of LTF-type, using recent rainfall and observed discharge data, (vii) n-AR, an AR model applied to the observed discharge series only, as a naïve updating model; and (viii) n-NARXM, a naïve form of the NARXM, using only the observed discharge data, excluding exogenous inputs. The five GFMFS simulation (non-updating) models used were the non-parametric and parametric forms of the Simple Linear Model and of the Linear Perturbation Model, the Linearly-Varying Gain Factor Model, the Artificial Neural Network Model, and the conceptual Soil Moisture Accounting and Routing (SMAR) model. As the SMAR model performance was found to be the best among these models, in terms of the Nash-Sutcliffe R^2 value, both in calibration and in verification, the simulated outflows of this model only were selected for the subsequent exercise of producing updated discharge forecasts. All the eight forms of updating models for producing lead-time discharge forecasts were found to be capable of producing relatively good lead-1 (1-day ahead) forecasts, with R^2 values almost 90% or above. However, for higher lead time forecasts, only three updating models, viz., NARXM, LTF, and NNU, were found to be suitable, with lead-6 values of R^2 about 90% or higher. Graphical comparisons were made of the lead-time forecasts for the two largest floods, one in the calibration period and the other in the verification period.

Keywords: forecast updating, autoregressive model, linear transfer function, neural networks

Introduction and objective of the study

The main purpose of river flow forecasting systems is the forecasting of flood magnitudes in real time so as to give timely warning to the water management authorities and other end-users of an impending flood at a particular location or to estimate the progress of a flood wave in a critical or 'high-alert' scenario. The lead-times of such forecasts generally vary from six hours to six days or more, depending on catchment size, etc. Apart from the hydrological modelling considerations, the accuracy of such forecasts is clearly dependent on the quality of the quantitative precipitation forecasts (*QPFs*) available which, for lead times of more than a few days, drops off quite rapidly. Time series analysis techniques applied for precipitation forecasting can be disappointing, particularly for the daily time step, but in the hourly time step case when coupled with a rainfall-runoff model which is itself coupled to a discharge error forecasting model, it can provide real improvement in the forecasts (Brath *et al.*, 2002). In their study, univariate time series models, both linear (ARMA) and non-linear (Neural Networks and Nearest Neighbour) models were used to improve the forecasts of a conceptual model, both for *QPF* estimation and for discharge updating. Hence meteorologists hold a vital key to successful river flow forecasting and real progress in such forecasting depends on the degree to which meteorologists become full and equal partners in the development of flow forecasting systems. Apart from flood forecasting, such flow forecasting systems, when operating in non-updating (simulation) mode, (using the exogenous inputs of precipitation and possibly also the recent outputs generated by the model but excluding the use of the recently observed discharges as model inputs), are of use in the day-to-day monitoring, operation, control, and management of water resource projects and hydraulic structures. In contrast to the 'Rainfall to Runoff' transformation models which operate without forecast updating, in 'simulation (i.e. design)' mode, real-time flood forecasting models and procedures attempt to compensate for the deficiency in matching the simulation mode hydrograph to the corresponding observed hydrograph. This is achieved by utilising additional input information from the most recently measured discharges. The quantities that may be updated are the discharge, the parameters, the state variables, or the input (WMO, 1992). Only discharge updating is considered in this study.

Designing a river flow forecasting system involves: understanding the behaviour of the catchment response to rainfall, devising or selecting a suitable rainfall–runoff model, incorporating meteorological forecast information for use in real-time flow forecasting, and producing regular reliable forecasts, over the required range of lead-times, at the necessary time intervals. A rainfall–runoff model can be selected from any of the three broad categories (see O'Connell, 1991), *viz*, empirical metric or 'black-box' models, 'conceptual' quasi-physical models, and grid (or pixel)-type 'distributed physical-process' models. Wheater *et al.* (1993) emphasised hybrid metric-conceptual (HMC) models as a separate model category, and Young (2002) showed that transfer-function-type models can efficiently characterise the relations between the rainfall-runoff data. Statistically, such data are amenable to simple conceptual interpretation (as HMC models), and are also ideal for incorporation in a real-time adaptive forecasting system based on recursive state-space estimation having the form of an adaptive Kalman filter. Whether on-line recalibration of hydrological model parameters at each time step is warranted is debatable (WMO. 1992). Other types of hybrid models such as the HL-RMS, developed by the US National Weather Service (Koren *et al.*, 2003, 2004), attempt to bridge the divide between the distributed physically-based and the lumped conceptual model. Combining these two approaches, transforms a good conceptual model into something more akin to a physically-based model, in the sense that its parameters can be estimated from physical properties rather than by calibration. Similarly, the new distributed hybrid TOPKAPI rainfall–runoff model (Ciarapica and Todini, 2002; Liu and Todini, 2002), which evolved from consideration of the conceptual Xinanjiang-ARNO model and the 'physically-conceived' TOPMODEL, "appears to be a promising tool worthy of further investigation".

Depending on the model structure, a discharge forecast updating facility may be an integral part of the model, as in a simple linear transfer function model (e.g. the PSLM and PLPM models), but it is more often an 'add-on' component that is operated in sequence or in parallel with the substantive model and calibrated separately. While both of these updating approaches are represented in the present study, the emphasis here is on discharge forecast updating as a separate issue to that of modelling the rainfall–runoff relation (with no consideration of recursive parameter updating or of state updating by Kalman filter, etc.) and on the comparison of the performances of all of the eight updating procedures now incorporated in the **G**alway River **F**low **M**odelling and **F**orecasting **S**ystem (GFMFS), as the forecast lead time is increased.

Lying at the bottom of the scientific scale, the empirical input-output 'black-box' (system/theoretic, or metric) models, generally 'lumped' or 'semi-lumped', simply attempt to relate precipitation input to stream flow as output, with little or no regard to the individual hydrological processes involved. It is largely an exercise in pattern recognition, selection of a structure to mimic the essentials of the pattern and curve fitting or calibration of the model. Neural networks are examples of versatile (but not parsimonious) 'black-box' models.

The physically-inspired 'conceptual' models, on the other hand, usually lumped but also amenable for use in 'semi-lumped' and in 'semi-distributed' form, attempt to simulate the perceived dominant hydrological mechanisms of the catchment response to rainfall, e.g. interception, evapotranspiration, infiltration, snowmelt, and both groundwater and surface water flow routing, perhaps including interflow, etc., using simple prescribed physically-plausible empirical and heuristic mathematical relations for each mechanism represented.

In contrast, the more sophisticated distributed physically-based models which, by definition, are 'distributed' both as regards inputs and hydrological processes, are based explicitly on current understanding of the physics of the hydrological processes involved in the generation of run-off and on accounting for the areal distribution, not only of the rainfall but also of hydrological mechanisms and variables such as storage. While the 'bottom-up' distributed physically-based models are undoubtedly far more scientific than the other three 'top-down' model types (including the HCM models), and have dominated modelling research for more than a decade now, their application in operational river flow forecasting systems is not yet widespread. This is partly due to their inherent complexity, substantial software and training costs, extensive data demands and still unresolved problems of scale in applying the physical laws for point processes to the grid elements or pixels of a digital elevation model. However, it is also due to a lingering perception among operational flood forecasters and some researchers that the river flow forecasts produced by such complex 'bottom-up' models are not necessarily superior to those of simpler, more primitive top-down' models and they can indeed be worse (Michaud and Sorooshian, 1994; Seyfried and Wilcox, 1995; Woolhiser, 1996; Ye *et al.*, 1997; Smith *et al.*, 2003; Reed *et al.*, 2004). However, both camps have their enthusiastic supporters.

Evidence of recent substantial progress in distributed modelling is the emergence of the spatially distributed water balance model (LISFLOOD), the multi-purpose modelling tool that has been developed explicitly for the simulation of floods in large European drainage basins. It is an element of the ambitious European Flood Forecasting System (EFFS) Project (2000–2003), funded by the EC-Fifth Framework Programme, which has the objective of combining state-of-the-art expertise in meteorology and hydrology on a European scale, the ultimate aim being to issue a 10-day pre-warning of floods. While there is general agreement that real and sustained progress is being made in distributed modelling (Ciarapica and Todini, 2002; Liu and Todini, 2002; Koren *et al.*, 2003), there is still a lack of consensus on the final outcome of such efforts, scale and parameter estimation still being identified as the biggest hurdles to overcome in distributed modelling (Koren *et al.*, 2004). No 'distributed physically-based' models are used in this study, but some of the updating procedures considered could be used just as effectively with such models as with those lacking in scientific pedigree. While the 'black-box' lineage of these relatively simple updating procedures might well render them anathema to developers of 'physically-based' models, operational forecasters might well adopt a more pragmatic and less fundamentalist approach!

Apart from naïve flow forecasting models of the time-series variety, based solely on the observed discharge series, updating information may be incorporated in a 'real-time flood forecasting system' in one of three broad categories:

(i) the rainfall-runoff model is first calibrated and run in simulation mode and subsequently estimated corrections, based on a separately calibrated empirical time-series model of previously observed forecast errors, or of the corresponding simulated and observed discharges, are applied to the simulation mode output forecast;
(ii) the rainfall-runoff model and the corresponding *'error correction'* or updating component are calibrated simultaneously, offline, using previously observed data, and both are applied simultaneously, in an integrated manner, for real time forecasting, and
(iii) an adaptive modelling structure is applied, whereby the model parameters and/or states are updated recursively, perhaps even at each time step, e.g. by recursive least-squares (RLS), the instrumental variable method (Young 1974; 1984; Young and Jakeman, 1979), or automatic updating through the use of an extended Kalman filter that provides the capability for real-time probabilistic forecasts of flood occurrence and flood magnitude (Georgakakos, 1986), a recent example of which, in the context of transfer-function-type models, is the 'state estimation' form applied by Young (2002) as a state estimation and data assimilation device.

Only the first two of the three categories listed above, however, are represented in this study of updating flow forecast procedures.

The specific objective of the present study is to evaluate the flow forecast performance of the eight updating models and procedures that have been incorporated in the GFMFS (O'Connor *et al*., 2001; Goswami *et al.,* 2002a,b) and to compare their resulting forecast efficiencies in lead-time forecasting in the context of an Irish catchment, for lead-times of one to six days. As the chosen updating models encompass a reasonably wide range of procedures, from simple Auto-Regressive models to the more complex Neural Networks, other well-known updating procedures or models, albeit more sophisticated and statistically superior (e.g. Young, 1974, 1984), were not included. Such a comparative study has been carried out for the River Brosna catchment in Ireland, for the Ferbane gauging station, (at which the contributing catchment area is 1207 km^2), using *daily* rainfall, evaporation and discharge data, for the period 1996 to 2001 inclusive.

The GFMFS software package used in the study

The GFMFS is a Windows-based software package, recently developed in the Department of Engineering Hydrology, National University of Ireland, Galway, having evolved from a series of seven International Workshops on River Flow Forecasting, held in Galway between 1985 and 1997. Extended for this study, it now comprises a suite of models for simulation, updating and real-time flow forecasting application, for lead-times of 1 to 6 days. Although reliable quantitative precipitation forecasts (QPFs) of more than a few days ahead are not currently available, discharge forecasts of up to six days ahead are used in this study for the purpose of discriminating between the effectiveness of the updating models, using actual rainfall data rather than QPFs as input, as explained below.

The application of these models in real-time forecasting mode, in common with other hydrological forecasting models, requires knowledge (or at least the best possible estimates available) of the input values over the lead-time of the discharge forecast. In the present work, as in other heuristic hydrological research studies (e.g. WMO, 1992; Kachroo and Liang, 1992), where the models are tested on the historical data but mimic the 'real-time' operational mode, the 'perfect input foresight over the lead-time of the output forecast' scenario is used as an operational input scenario. The adoption of this 'ideal' input scenario, apart from the simplification it introduces, at least eliminates those errors and uncertainties introduced by imperfect knowledge of the input variables over the forecast lead-time, i.e. the deficiencies in the QPFs of the precipitation forecasting model, so that the intercomparison of the discharge updating models is unaffected by complications arising from the efficiency or otherwise of that QPF model. Thus, while the form of updating considered in this study is 'pseudo-real-time', rather than 'real-time', the updating models can equally well be applied in the 'real-time' updating scenario but with a much lower forecasting efficiency for the higher forecast lead-times.To assess (and distinguish between) the performance of the eight updating models, it is necessary to use forecast lead-times of up to six days.

The GFMFS reflects a small fraction of the modelling efforts of the Galway International Postgraduate Hydrology Courses (now sadly terminated). It incorporates five of the simpler rainfall-runoff models routinely used in Galway (four system/theoretic 'black-box' and one 'conceptual') to be run in simulation (non-updating) mode and eight models (including two naïve models) to be run in updating mode. The GFMFS is also capable of combining either non-updated or updated flow forecasts by the methods of simple average (SAM), weighted average (WAM) and artificial neural network (ANN), to produce 'consensus' forecasts (Shamseldin *et al.*, 1997). Descriptions of the models and combination techniques incorporated in the GFMFS are widely available, e.g. in O'Connell *et al.* (1970); Nash and Foley (1982); Nash and Barsi (1983); Khan (1986); Kachroo *et al.* (1988); Kachroo (1992, a, b); Liang (1992); Ahsan and O'Connor (1994); Liang *et al.* (1994); Zhang *et al.* (1994); Tan and O'Connor (1996); Shamseldin (1997); Shamseldin *et al.* (1997); Xiong *et al.* (2001). Brief descriptions of these simulation (non-updating) models are given in the Appendix.

The forecast updating models

Eight forecast-updating models were applied in this study. Persistence (e.g. a tendency for high values to follow high values, as displayed by its sample autocorrelation function) is a recognised characteristic of observed discharge series. This arises primarily from the storage effects of the system but also from some degree of persistence in the exogenous system inputs such as rainfall. Seasonal influences and, to a lesser extent, some element of long-term trend due to changes in agricultural practices, climate change, urbanisation, etc. may also be present in the discharge series. Likewise, persistence in the values of model output residuals (i.e. model forecast errors) of both conceptual and black-box models is also common. In the case of a model-error series, rendered stationary if necessary by differencing or some other transformation, persistence is taken to be an indication that not all of the deterministic nature of the input-to-output relation was captured by the model, as otherwise the errors might be expected to consist of pure 'white noise' rather than being 'coloured' by a persistence structure. Therefore, in operational real-time forecasting, modellers conventionally attempt to exploit such persistence by using a univariate error-forecasting model designed to estimate the errors likely to occur in the immediate future, i.e. over the next few time steps. These error forecasts are then added as a correction to the corresponding forecasts of the substantive model to provide the updated forecasts. Such coupling of a substantive model with its corrective error-forecasting model is perhaps the most widely used forecast updating procedure but there are other forecasting options, some of which are 'causal' (i.e. involving one or more exogenous input variables) rather than 'univariate' in approach.

Whereas persistence in the error series of a rainfall-runoff model can be used to advantage in forecast updating, heteroscedasticity in the time-series of errors (i.e. variability in the variance), which is also a recognised characteristic of such series, is an unwelcome complication in least-squares estimation. However, the assumption of homogeneity of variance, i.e. homoscedasticity, greatly simplifies mathematical and computational treatment and may lead to good (but admittedly statistically inferior) estimation results even if the assumption is not true. In the updating procedures applied in this study, heteroscedasticity in the time-series of errors was not taken into account.

The updating procedures applied in this study, using the

Brosna data, were:

(i) the standard linear Auto-Regressive (AR) model, applied to the residuals obtained after applying a substantive simulation (non-updating) rainfall-runoff model to the catchment data;

(ii) the Neural Network Updating (NNU) model, also applied to such residuals as input;

(iii) the Linear Transfer Function (LTF) model, applied to the simulated and the recently observed discharges;

(iv), the Non-linear Auto-Regressive eXogenous-Input Model (NARXM), also a neural network-type structure, but having wide options of using recently observed values of one or more of the three Brosna data series, together with the non-updated simulated outflows over the forecast lead-time, as inputs;

(v) the Parametric Simple Linear Model (PSLM), of the LTF-type, using recent rainfall and discharge observations;

(vi) the Parametric Linear perturbation Model (PLPM), also of the LTF-type, using recent rainfall and discharge observations;

(vii) n-AR, an AR model applied to the observed discharge series only, as a naïve updating model; and

(viii) n-NARXM, a naïve form of the NARXM, using only the observations of discharge, excluding all exogenous inputs.

These eight forecast updating models clearly fall into the 'black-box' category, and all are used for short lead-time forecasts (of days rather than weeks).

THE AUTO-REGRESSIVE (AR) FORECAST ERROR ESTIMATION MODEL FOR FORECAST UPDATING

The classic model of the persistence structure of a time series is the simple linear univariate Auto-Regressive (AR) model. As the persistence structure of the output of a linear storage system subjected to a 'white-noise' input (which by definition has no persistence) is identical to the persistence structure of the unit impulse response of the system, the objective of AR modelling is to identify an AR model structure such that the persistence structure of its unit impulse response series matches, as closely as possible, the persistence structure of the time series being modelled, as reflected by its serial autocorrelation function. If such a model can be identified and its input (obtained by back-routing the time series through the model) shown by its correlogram to be essentially 'white noise', then the model can be considered a good approximation of the generating mechanism of that time series. Mathematically, an AR process of order p is defined by the equation (Box and Jenkins, 1976)

$$e_{t+1} = \bar{e} + \phi_1(e_t - \bar{e}) + \phi_2(e_{t-1} - \bar{e}) + \ldots\ldots + \phi_p(e_{t-p+1} - \bar{e}) + a_{t+1} \quad (1)$$

in which e_t is the model forecast error (or 'residual') at time t, the series having mean $\bar{e}$, a_t is the value at time t of a pure 'white noise' sequence with zero mean and constant variance, σ_a^2, and the ϕ_i, for $i = 1$ *to* p, are the parameters of the auto-regressive model. The residual e_t in the AR forecast-error updating model (Fig. 1) is defined by $e_t = Q_t - \hat{Q}_t$, for all values of t. The parameters ϕ_i of the AR model are usually estimated by the Yule-Walker (Box and Jenkins, 1976, pp. 82–84) or more simply by the method of ordinary least squares (OLS), the former method having been adopted in the case of the GFMFS.

Having calibrated the AR model, an estimate of the lead-1 (the *'one-step-ahead'*) forecast $\hat{e}_{t+1/t}$, made at the forecast time origin t, of the error e_{t+1} at time $t+1$, is given by

$$\hat{e}_{t+1/t} - \bar{e} = \hat{\phi}_1(e_t - \bar{e}) + \hat{\phi}_2(e_{t-1} - \bar{e}) + \ldots\ldots + \hat{\phi}_p(e_{t-p+1} - \bar{e}) \quad (2)$$

on taking the expectation $E[a_t] = 0$. The *lead-L* forecast error $\hat{e}_{t+L/t}$, made from a forecast origin t, may be obtained from the following equation:

$$\hat{e}_{t+L/t} = \bar{e} + \hat{\phi}_1([e_{t+L-1}] - \bar{e}) + \hat{\phi}_2([e_{t+L-2}] - \bar{e}) + \ldots\ldots + \hat{\phi}_p([e_{t+L-p}] - \bar{e}) \quad (3)$$

where the square brackets denote the conditional expectation of the quantity between the brackets.

Having obtained the estimate of the discharge forecast error for the desired lead time $L \geq 1$, the updated discharge forecast for that lead time, for a forecast time origin t, is given by

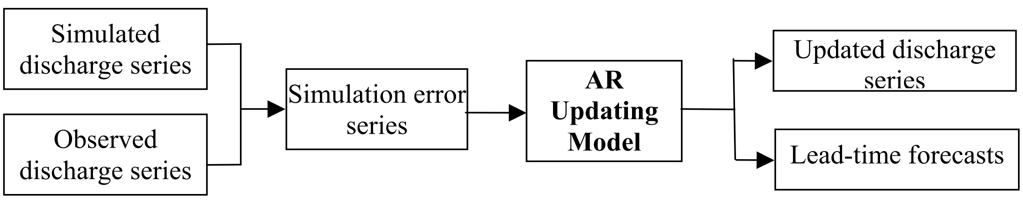

Fig. 1. *Schematic diagram of the standard linear Auto-Regressive (AR) updating model*

$$\hat{\hat{Q}}_{t+L/t} = \hat{Q}_{t+L} + \hat{e}_{t+L/t} \tag{4}$$

where $\hat{Q}_{t+L/t}$ is the corresponding simulation forecast of the substantive rainfall-runoff model (based on recently observed inputs and on forecasted inputs over the lead-time L) and $\hat{e}_{t+L/t}$ is the estimate of its error.

Thus, having estimated the forecast errors $\hat{e}_{t+L/t}$ for specified lead times L using the auto-regressive procedure, the corresponding updated forecasts of the outflow can be obtained by applying the errors successively to the corresponding estimates $\hat{Q}_{t+L}$ of the outflows for those lead-times obtained by applying the substantive model. It may be noted that the forecast values of inflows over the forecast lead-time L must first be used to obtain these non-updated $\hat{Q}_{t+L}$ estimates. For regular real-time flow forecasting (for the selected time step), the procedure for issuing the updated forecasts for the specified lead-times has to be repeated for each time step, as soon as the newly observed flow value for that lead-*1* becomes available, thereby adding that lead-*1* value to the historical discharge record. This results in recasting the forecasts for the whole specified range of lead-times (1, L), at each time-step, i.e. for each successive forecast origin.

THE LINEAR TRANSFER FUNCTION (LTF) MODEL FOR FORECAST UPDATING

Recognising that the AR forecast error estimation model of the last section can be written as

$$\Phi(B)(e_t - \bar{e}) = \Phi(B)\left\{(Q_t - \overline{Q}) - \left(\hat{Q}_t - \overline{\hat{Q}}\right)\right\} = \Phi(B)(Q_t - \overline{Q}) - \Phi(B)\left(\hat{Q}_t - \overline{\hat{Q}}\right) = a_t$$

Peetanonchai (1995), as a generalisation of AR model, suggested the linear transfer function LTF form

$$A(B)(Q_t - \overline{Q}) - C(B)\left(\hat{Q}_t - \overline{\hat{Q}}\right) = a_t \tag{5}$$

for use in forecast updating, in which $A(B) = \left[1 - a_1 B - a_2 B^2 - .... - a_p B^p\right]$ and $C(B) = \left[c_0 + c_1 B + c_2 B^2 + ... + c_q B^q\right]$, a concept extended to real-time 'consensus' forecasting by Shamseldin and O'Connor (1999).

Since $A(B)\overline{Q} = \left[A(B)\right]_{B=1} \times \overline{Q}$ and $C(B)\overline{\hat{Q}} = \left[C(B)\right]_{B=1} \times \overline{\hat{Q}}$, and recognising that the 'Gain Factor' G of the LTF model has the form $G = \left[\frac{C(B)}{A(B)}\right]_{B=1} = \frac{\overline{Q}}{\overline{\hat{Q}}}$ the LTF updating model defined by Eqn.(5) reduces to the standard form applied in the GFMFS for this study, i.e.

$$A(B)Q_t = C(B)\hat{Q}_t + a_t \tag{6}$$

in which the parameters of the model are estimated by the method of **O**rdinary **L**east **S**quares (*OLS*). While the *OLS* parameter estimates of the LTF model are expected to be asymptotically biased, and recursive procedures which produce unbiased estimates (Ljung and Soderstrom, 1983; Young, 1984; Norton, 1989) would be better, this problem was not addressed as it was not considered to be essential in the context of comparing the lead-time discharge forecasts of the selected eight updating models, using daily data.

Having calibrated the model, its updated Lead-1 forecast of the discharge, for a forecast time origin t, is given by

$$\hat{Q}_{t+1/t} = \left[1 - \hat{A}(B)\right]Q_t + \hat{C}(B)\hat{Q}_{t+1/t} = \left[\hat{a}_1 B + \hat{a}_2 B^2 + ... + \hat{a}_p B^p\right]Q_{t+1} + \\ + \left[\hat{c}_0 + \hat{c}_1 B + \hat{c}_2 B^2 + ... + \hat{c}_q B^q\right]\hat{Q}_{t+1} \tag{7}$$

B being the backward shift operator, the Lead-2 forecast being

$$\hat{Q}_{t+2/t} = \hat{a}_1 \hat{Q}_{t+1/t} + \hat{a}_2 Q_t + ... + \hat{a}_p Q_{t-p+1} + \hat{c}_0 \hat{Q}_{t+1} + \hat{c}_1 \hat{Q}_t + \hat{c}_2 \hat{Q}_{t-1} + \\ = \hat{a}_1 Q_t + \hat{a}_2 Q_{t-1} + ... + \hat{a}_p Q_{t-p+1} + \hat{c}_0 \hat{Q}_{t+1} + \hat{c}_1 \hat{Q}_t + \hat{c}_2 \hat{Q}_{t-1} + ... + \hat{c}_q \hat{Q}_{t-q+1} \\ - .... + \hat{c}_q \hat{Q}_{t-q+1} \tag{8}$$

and so on. Thus, a linear transfer function model is applied as a generalisation of the AR forecast error estimation method of the previous section. Note that the LTF model collapses back to the AR form if $A(B) = C(B) = \Phi(B)$. A schematic diagram of an LTF updating model is given in Fig. 2. Despite their simplicity, AR models are still effective forecast error estimation tools (Xiong *et al.*, 2001).

THE NEURAL NETWORK UPDATING (NNU) MODEL

This artificial neural network is applied, in the GFMFS package, in the context of providing a non-linear function mapping of a set of inputs (the inputs being the errors of forecast estimation based on the flow values produced by the substantive model operating in simulation mode) into the network output (i.e. the updated lead-time forecast discharge). However, the specific mathematical form of the relationship is unspecified, although the network structure must be pre-configured before training. The structure of the neural network, as used in updating mode, is similar to that described in the appendix and shown in Fig. A.3.

THE NON-LINEAR AUTO-REGRESSIVE EXOGENEOUS-INPUT MODEL (NARXM)

In this updating procedure, another neural network is applied, in what may be considered loosely as a non-linear auto-regressive formulation, providing a non-linear function

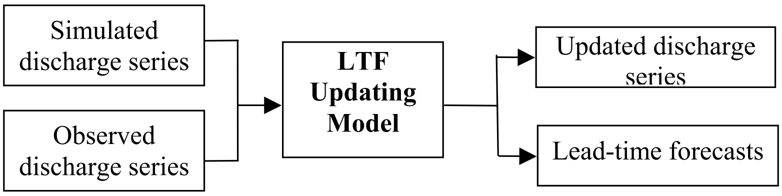

Fig. 2. *Schematic diagram of the Linear Transfer Function (LTF) updating model*

mapping of a set of exogenous inputs (i.e. the input rainfall, input evaporation, the substantive model estimated outflows over the lead time, and the most recent observed outflows) into the single network output (i.e. the updated lead-time forecast of the discharge) (Shamseldin and O'Connor, 2001). The structure of the neural network is similar to that given in the appendix and shown in Fig. A.3.

THE LTF-TYPE PARAMETRIC SIMPLE LINEAR MODEL (LTF-PSLM)

In this base-line LTF-type parametric simple linear model (LTF-PSLM), the representation of the transformation process of the input rainfall series R to the output discharge series Q, for discrete time steps, is given by

$$Q_t = \sum_{j=1}^{r} \alpha_j Q_{t-j} + \sum_{j=0}^{s} \omega_j R_{t-b-j} + e_t \tag{9}$$

i.e. the R series is used as part of the exogenous input rather than the simulated outputs $\hat{Q}$ used in the LTF updating model described earlier. The model is calibrated by the method of ordinary least squares (*OLS*), (ignoring the effects of the resulting parameter estimates being asymptotically biased), the updating form of the model, for a forecast time origin t, being

$$\hat{Q}_{t+1/t} = \sum_{j=1}^{r} \hat{\alpha}_j Q_{t-j+1} + \sum_{j=0}^{s} \hat{\omega}_j R_{t-b-j+1} \tag{10}$$

When operating in the non-updating mode, the past computed values of Q are used on the right- hand side of the transfer function equation, whereas in updating mode, as considered here, the most recent observed values of Q are to be used (on the right-hand side) as input to the model. While the LTF-PSLM is clearly 'naïve' in the context of rainfall-runoff modelling, it is not so in the context of updating models, where 'naïve' refers to an updating model based solely on the discharge record.

THE LTF-TYPE PARAMETRIC LINEAR PERTURBATION MODEL (LTF-PLPM)

The mathematical form of the LTF-type parametric linear perturbation model (LTF-PLPM) for updating is identical to that of the LTF-PSLM, the only difference being that the series of departures of inputs and outputs from their seasonal mean values are used in the transfer function equation instead of the actual recorded series as used in LTF-PSLM. The updated output departure forecast is then simply added to the corresponding seasonal forecast to give the updated discharge forecast of the LTF-PLPM.

THE NAÏVE AUTO-REGRESSIVE (n-AR) MODEL

A naïve linear auto-regressive updating model may be fitted to the observed discharge series alone provided it displays significant persistence. For order p, the univariate n-AR model has the form,

$$\hat{Q}_{t+l/t} = \overline{Q} + \hat{\phi}_1\left(\left[Q_{t+l-1}\right] - \overline{Q}\right) + \hat{\phi}_2\left(\left[Q_{t+l-2}\right] - \overline{Q}\right) + \ldots + \hat{\phi}_p\left(\left[Q_{t+l-p}\right] - \overline{Q}\right) \tag{11}$$

i.e. using the observed discharge series rather than the forecast error series used in the AR forecast-error estimation model described earlier. The same principle as used for forecast error estimation is used for such naïve estimation of the discharge forecast. The updated forecast flow values for the specified lead-times can be estimated directly by iteratively using the auto-regressive equation, once the model parameter values have been estimated. Hence, it is a simple univariate time series forecasting model, considered as a naïve ('base-line') model against which the performance of more substantive updating models, incorporating the simulated forecasts of a rainfall–runoff model, can be compared. It is not suitable for forecasting more than one or two steps ahead.

THE NAÏVE NON-LINEAR AUTO-REGRESSIVE EXOGENEOUS-INPUT (n-NARXM) MODEL

The principle of the naïve form of non-linear auto-regressive exogenous-input (n-NARXM) updating model, having the structure of a neural network, is similar to that of the naïve auto-regressive updating model, but in this case the lead-time forecast of outflows is expressed as a finite, non-linear aggregate of current and previous discharge values, instead of the linear weighting in the case of the n-AR model. As it is much more general than the naïve AR updating model, one would reasonably expect it to perform better. Both are 'black-box' univariate updating models.

The model efficiency evaluation criteria

The performance of a model must be judged on the extent to which (i) it satisfies its objective of simulating the real world phenomenon (accuracy), (ii) the achieved level of accuracy persists through different samples of data (consistency) and (iii) it can sustain the achieved level of accuracy when subjected to diverse applications and tests other than those used for calibrating the model (versatility). A forecast efficiency criterion is therefore necessary to judge the performance of a model. Fifteen different indices of model performance evaluation are included in the GFMFS package, viz., (i) Mean Square Error (*MSE*), (ii) Normalised Error Function Value, (iii) Coefficient of Efficiency (R^2), (iv) Modified Index of Efficiency R^2_j (v) Index of Volumetric Fit (*IVF*), (vi) Index of Agreement, (vii) Modified Index of Agreement, (viii) Coefficient of Determination, (ix) Relative Error (RE) of the Peak, (x) Root Mean Square Error (RMSE), (xi) Square-Root-RMSE, (xii) % Bias, (xiii) Log RMSE, (xiv) Square-Root-Nash-Sutcliffe Efficiency, and (xv) Visual Comparison."

In the present study, the Nash-Sutcliffe (1970) model performance evaluation index, R^2, which is based on the Mean Square Error (*MSE)*, is adopted as the primary model efficiency index. While it is used, in the GFMFS, as the main index, the visual comparison method is also employed on selected events. Other expressions of model forecast error include the Mean Absolute Deviation (*MAD*) but the *MSE*, which penalises the forecasting model much more for large errors than for small errors, irrespective of the magnitude of the variable at which such errors occur, is far more widely used by catchment modellers and it forms the basis of the objective function used to calibrate the models in the *GFMFS*. Clearly, both R^2 and *MSE* are global indices, representative of the entire period under consideration. Although the R^2 index is convenient (many model calibration methods are based on a least-squares matching of model output to observed output), its relative insensitivity is also well recognised and its weakness, particularly when the mean of the calibration differs greatly from that of the validation period, e.g. one wet and the other dry, is well known (Kachroo and Natale, 1992). It remains, however, the most widely used model efficiency index.

Note that the type of forecast considered here is the 'point' forecast, i.e. a single number representing the 'best' prediction, rather than the 'prediction interval' type of forecast. Increasingly, as the need to provide a rigorous framework for the uncertainties associated with model parameters and forecasts is addressed, and the need to link the forecasts to an assessment of the associated hazard risks and costs receives more attention, the emphasis has perhaps shifted away from trying to improve the 'point' forecast. However, as the authors consider that the 'point' forecast is still important, it is the focus of this study.

Modelling in simulation and updating modes and discussion of the results

All the five substantive models in simulation mode were applied individually to estimate the discharge from the given rainfall and evaporation data of the River Brosna catchment (up to the Ferbane gauging station). Out of the total of 2192 (1996–2001) days of daily discharge data, the weighted average daily rainfall and daily evaporation data, the first 1461 days of data were used for model calibration and the last 731 days for verification. Table 1 shows the values of the Nash and Sutcliffe (1970) forecast performance evaluation index, R^2, for each of the substantive models, all operating in simulation mode. It is observed, from Table 1, that the non-updated forecast performance of the SMAR model is clearly the best, in terms of R^2, having values of 85.0% and 78.17% for the calibration and verification periods respectively.

In applying the first type of updating noted in the 'Introduction', whereby the rainfall-runoff model is first calibrated and run in simulation mode and subsequently correction estimates (based on a separately calibrated empirical time-series model) are applied to the simulation mode output forecasts to obtain updated forecasts. In this study, the simulated output from the substantive SMAR model was used along with the observed discharge data to obtain the error series for application in the AR and the NNU updating models. The most recent simulated discharges of the SMAR model were also used as exogenous inputs in the LTF and NARXM updating models. However, for the other four updating models, namely P-SLM, P-LPM, n-AR and n-NARXM which involve the second type of updating noted in the Introduction, the coupled rainfall-runoff model and updating facility are calibrated simultaneously, offline, using previously observed data, and are likewise subsequently applied simultaneously (as a coupled model) to produce the forecasts, the simulated output from the SMAR model is not involved. These six updating models, together with the naïve n-AR and the n-NARXM, were used to estimate the updated forecasts for the forecast lead times of 1, 2, 3, 4, 5 and 6 days.

The values of R^2 in the calibration and verification periods

Table 1. Values of Nash-Sutcliffe efficiency index, R^2, for each model, in simulation mode

S.N.		*Model*	*Values of R^2 in %*	
			Calibration	*Verification*
1	a	Non parametric simple linear model (NP-SLM)	40.67	39.69
	b	Parametric simple linear model (P-SLM)	33.87	31.56
2	a	Non parametric linear perturbation model (NP-LPM)	76.12	75.70
	b	Parametric linear perturbation model (P-LPM)	78.53	71.92
3		Linearly varying gain factor model (LVGFM)	44.85	46.62
4		Artificial neural network model (ANN)	73.73	60.65
5		Soil Moisture Accounting and Routing model (SMAR)	85.00	78.17

respectively, for each of the updating models, are presented in brackets in Table 2. This table shows the ranking of the updating models, in terms of R^2 values in calibration (R_c^2), for the lead-1 to lead-6 day forecasts. It is seen in Table 2 that, for the lead-1 forecast, among the eight updating models considered in the present study, two models, viz. the non-linear auto-regressive exogenous-input neural network model (NARXM) and the linear transfer function (LTF) forecast model gave the best forecasts, with values of R_c^2 exceeding 95%. Five other updating models, viz. the LTF-type parametric simple linear model (P-SLM), the LTF-type parametric linear perturbation model (P-PLPM) model, the Neural Network Updating (NNU) forecast error model, the standard linear AR-error model and the naïve NARXM also gave good forecasts, with values of R_c^2 above 90%. Even the eighth updating model, the naïve AR-model, gave a relatively good forecast, its value of R_c^2 being quite close to 90%. To summarise, all eight updating models considered are capable of producing relatively good lead-1 day forecasts.

For the lead-2 and lead-3 day forecasts, only three updating models, the non-linear NARXM, the linear transfer function model (LTF) and the NNU error updating model, gave very good forecasts, with values of R_c^2 higher than 90%. For the lead-4 to lead-6 days forecasts, only two models, the non-linear NARXM and the LTF updating model gave consistently good forecasts, with values of R_c^2 above 90%. Indeed, the NNU model also gave relatively good lead-4 day to lead-6 day forecasts, the value of R_c^2 being quite close to 90%. Three other models, the two LTF-type updating models based on rainfall input and the AR-error updating model, gave fairly good higher lead-time forecasts, with values of R_c^2 = 76% or more. The remaining two naïve models, which do not use the rainfall input, have values of R_c^2 barely above 50%, showing quite clearly that all of the updating models based on rainfall perform significantly better than the two naïve models based on observed discharges only.

So, the results expressed solely in terms of R^2 indicate that the three models NARXM, the LTF, and the NNU are the most suitable for higher lead time forecasts (for lead-2 days and above). Even the other three models, P-SLM, the P-LPM and the AR-updating model, gave fairly good higher lead-time forecasts, whereas the two naïve models proved to be unsuitable, at least for the data set of the River Brosna catchment. The patterns for the R^2 values in verification are similar to those in calibration. As stated earlier, in all the above lead-time forecasting applications of the models, the 'perfect foresight of input over the forecast lead-time' scenario was used. Clearly, if the actual rainfall forecasts over the lead-times, for each day, were used instead of those of this scenario, the R^2 values would no doubt be lower than those achieved in this study, as the inaccuracies and uncertainties associated with the hydrological models would be compounded by those of the meteorological forecasts. Such rainfall forecasts were not available to the authors for this study.

As an extension of the above exercise, model output combination techniques were applied to each of the 6-lead day forecasts from the three best updating models NARXM, the LTF and the NNU model, so as to examine whether such 'consensus' forecasts would provide lead-time forecasts superior to those provided by the individual models. Values of R^2, for each of the 6-lead day combined forecasts, i.e. from the weighted average method (WAM) and the Neural Network combination method (NNAM), are shown at the bottom of Table 2. While there is some improvement in the value of R^2 for lead-1 day to lead-3 day forecasts, there is virtually no improvement in the value of R^2 for lead-4 to lead-6 day forecasts, i.e. in the case of the Brosna catchment, the 'consensus' updated forecasts for higher lead times were not significantly better than those of the individual models. Apart from consideration of the R^2 results, which only give a global picture of flow forecasting performance over the whole data set, the performance of the models on selected flood events was also examined. The lead-time forecasts

Table 2. Ranking of the 8 updating models for various forecast lead-times (in days), in terms of R^2: (R^2calibration, R^2verification)

Rank	*Lead -1*	*Lead - 2*	*Lead - 3*	*Lead -4*	*Lead - 5*	*Lead - 6*
1	NARXM (95.99, 96.41)	NARXM (94.01, 93.91)	NARXM (93.19, 92.97)	NARXM (92.35, 91.99)	NARXM (91.86, 91.49)	NARXM (91.63, 91.29)
2	LTF (95.79, 96.68)	LTF (93.51, 93.84)	LTF (92.58, 92.67)	LTF (91.69, 91.75)	LTF (91.15, 91.12)	LTF (90.77, 90.64)
3	P-SLM (94.45, 95.99)	NNU (90.81, 86.18)	NNU (90.13, 85.30)	NNU (89.64, 84.87)	NNU N(89.24, 84.48)	NNU (88.87, 84.20)
4	P-LPM (94.33, 96.09)	P-LPM (88.17, 91.18)	P-LPM (84.65, 88.68)	P-LPM (82.41, 86.67)	P-LPM (80.87, 84.68)	AR-model (81.10, 82.07)
5	NNU (92.12, 87.58)	P-SLM (88.08, 90.82)	P-SLM (83.89, 88.21)	AR-model (81.85, 83.62)	AR-model (80.76, 82.36)P-	P-LPM (79.72, 82.90)
6	AR-model (91.95, 94.01)	AR-model (84.38, 86.79)	AR-model (82.97, 85.40)	P-SLM (80.75, 85.95)	P-SLM (78.25, 83.76)	P-SLM (76.16, 81.61)
7	n-NARXM (90.15, 93.79)	n-NARXM (75.40, 85.25)	n-NARXM (66.60, 81.53)	n-NARXM (60.65, 77.91)	n-NARXM (55.60, 73.62)	n-NARXM (51.25, 69.61)
8	n-AR (89.75, 93.7)	n-AR (74.53, 85.20)	n-AR (65.29, 81.68)	n-AR (59.06, 78.30)	n-AR (54.35, 74.29)	n-AR (50.59, 70.50)
MOCT	WAM (96.10, 96.76) NNAM (96.09, 96.86)	NNAM (94.12, 93.88) WAM (94.09, 94.16)	NNAM (93.27, 93.02) WAM (93.22, 93.18)	WAM (92.34, 92.33) NNAM (92.14, 92.26)	WAM (91.84, 91.83) NNAM (91.61, 91.73)	WAM (91.56, 91.56) NNAM (91.53, 90.99)

Note:
NARXM - Non linear autoregressive eXogenous model
P-SLM - Parametric simple linear updating model
MOCT - Methods of output combination
LTF - Linear transfer function model
P-LPM - Parametric linear purturbation updating model
NNAM - Neural network averaging method
NNU - Nueral network updating model
n-NARXM - Naïve NARXM (Neural Network) model
WAM - Weighted averaging method
AR - Autoregressive updating model
n-AR - Naïve AR model

estimated by the 'best' three updating models, in terms of R^2, namely the NARXM, the LTF and the NNU, were compared with the corresponding observed hydrographs for the two highest floods in the record, i.e. the December 1999 (91.5 $m^3 s^{-1}$) flood, in the calibration period, and the November 2000 (85.65 $m^3 s^{-1}$) flood, in the verification period. The results are shown in Figs. 3 and 4. It is seen from Fig. 3a that all the three updating models are capable of producing lead-1 day forecasts that match the observed flows very well, both in terms of peak value and the time to peak, for this highest flood. Figures 3b and c show that the lead-2 and lead-3 day forecasts also match satisfactorily for these three '*best*' models. Figures 3d – 3f show that the lead-4 to lead-6 day forecasts produce a satisfactory match only in the case of the first ranked model, i.e. the NARXM updating model. In contrast, Figs. 4a and b, for the second highest flood (occurring in the verification period), show that only the LTF-updating model is capable of producing lead-1 day and lead-2 day forecasts that match the observed flows very well, whereas the forecasting performance of the NNU model for the lead-1 and -2 day forecasts is just about satisfactory. None of the lead-3 and -4 day forecasts of these three 'best' models produce a satisfactory match, but surprisingly the lead-5 and 6 day forecasts of the LTF and NNU updating models produced a satisfactory fit to the peak of the hydrograph. The NARXM updating model, which was ranked first in terms of the value of R^2 for both the calibration and verification periods, failed to reproduce the second largest flood well, for any of the lead times. These results serve as a reminder that the forecasts of any globally calibrated updating model, even if they give a high value of R^2, do not necessarily perform well on each individual flood event, particularly as regards the peak and time to peak.

Summary and conclusions

In all, five non-updating models and eight updating models, which are incorporated in the GFMFS software, were applied in this study, using precipitation, discharge and evaporation data from 1996–2001. The objective was to compare the updating model/procedures for generating the lead-1 to lead-6 day flow forecasts at the Ferbane gauging station of the River Brosna catchment in Ireland. Among the five non-updating models considered, the SMAR model was found to be the best, in terms of the Nash-Sutcliffe model forecast efficiency index R^2. Hence, for real-time forecasting, using the scenario of 'perfect foresight of input over the lead times', the SMAR-estimated non-updated discharge series was selected, for use with the observed discharges, to calibrate all the eight updating models available in the GFMFS. All eight were found to be capable of producing relatively good lead-1 forecasts, on the basis of the R^2 values, corresponding to almost R^2 = 90%. However, for higher (e.g. lead-2 to lead-6 days) lead times, only three models, the NARXM, the LTF and the NNU, produced good forecasts based on R^2 values. Three other models, the P-LPM, the P-SLM and the AR-error updating model, also gave fairly good higher lead-time forecasts. However, the two naïve models, the n-AR and the n-NARXM, which are based solely on the observed discharge, were unsuitable for higher lead time forecasting for the River Brosna at the Ferbane gauging station.

Zooming in on details of the forecast matching, graphical comparisons were made for the two highest floods, using the lead time forecasts of the three 'best' updating models NARXM, LTF and NNU. These show that the NARXM updating model, ranked first in terms of the value of R^2, is capable of producing very good lead-1 to -6 day forecasts for the highest flood occurring in the calibration period, but not so for the second highest flood occurring in the verification period. The LTF updating model, ranked second in terms of R^2, was capable of producing very good forecasts only for lead-1 day for the highest flood. In contrast, for the second highest flood, the first ranked updating model, i.e. the NARXM, was unable to produce good lead time forecasts, whereas the second ranked LTF-updating model produced very good lead-1 and lead-2 day forecasts for this flood.

Graphical comparisons of the results of some of the methods considered, in the context of forecasting the peak-flow and the time-to-peak, suggest that the non-adaptive updating procedures considered in this study are not satisfactory over the whole range of flood magnitudes, i.e. they have their successes but they also have their failures. This can be attributed partly to the global *MSE* objective function used to calibrate the models, as reflected in the corresponding global Nash-Sutcliffe R^2 forecast efficiency index employed to evaluate their performance. Clearly, if the forecasting of high flows, especially the peak-flow and the time-to-peak, is the sole purpose of the exercise, then an objective function which gives equal weight to a forecast error value, irrespective of the flow magnitude at which it occurs, is unlikely to provide a consistent optimum solution for the peak flows. A weighted least squares objective function (Zhang *et al.*, 1994), concentrating on the range of flow values of interest, might be a significant improvement. A detailed analysis of the distribution of forecast errors over the range of flow magnitudes might suggest a consistent pattern of inadequacy that could result in a change in model structure to compensate for such inadequacies in the performance of the models. Perhaps a global objective function such as the *MSE* is insufficient for peak flow forecasting and that a more sensitive and responsive local corrective procedure is required when things start to go badly wrong, i.e. some kind of objective but adaptive local calibration procedure may be necessary as a corrective measure. Some of these issues, relating to problems in peak flow forecasting, are currently under investigation by the authors.

Acknowledgements

The authors are pleased to express their sincere appreciation to Dr. Klara Finkele and Mr. Aidan Murphy, of Met Éireann, for providing the meteorological data, and to Ms. Clíona Ní Dhónaill, of the Irish Office of the Public Works, for providing the hydrological data. The first author gratefully acknowledges the support provided from the National University of Ireland, Galway, through its Millennium Research Fund, for his research on the development of the GFMFS package used in the present paper. The third author also gratefully acknowledges the research support provided by the Environmental Change Institute of the National University of Ireland, Galway, and by the Higher Education Authority of Ireland. The authors are greatly in debt to the two reviewers for their helpful comments; criticisms on the earlier version of this paper were deserved and most of them them have been addressed in the current version.

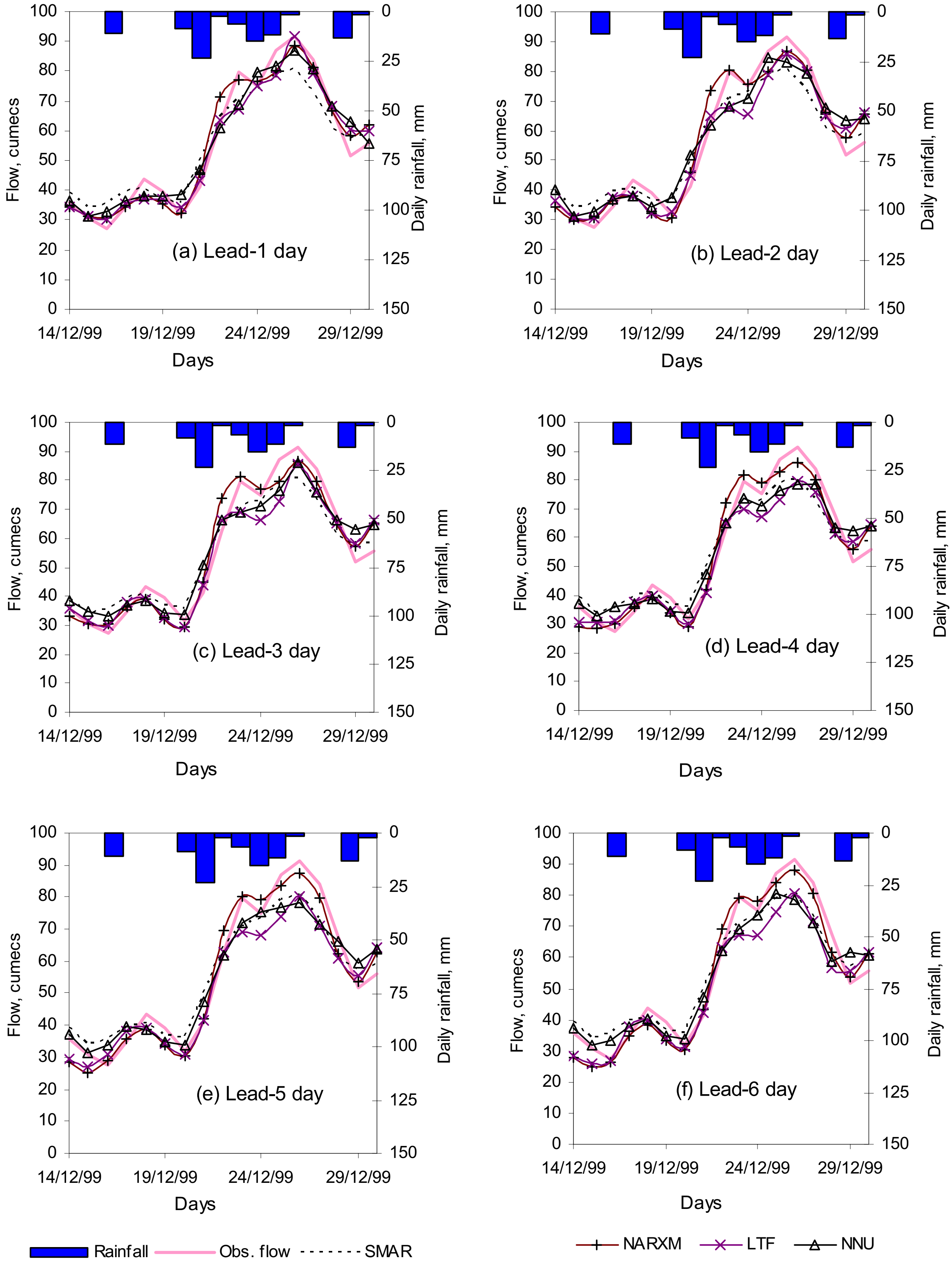

Fig. 3 (a-f). *Comparison of NARXM, LTF and NN-updating model forecasts for (a) lead-1 day to (f) lead-6 days, for the highest flood, of December 1999 (in the calibration period).*

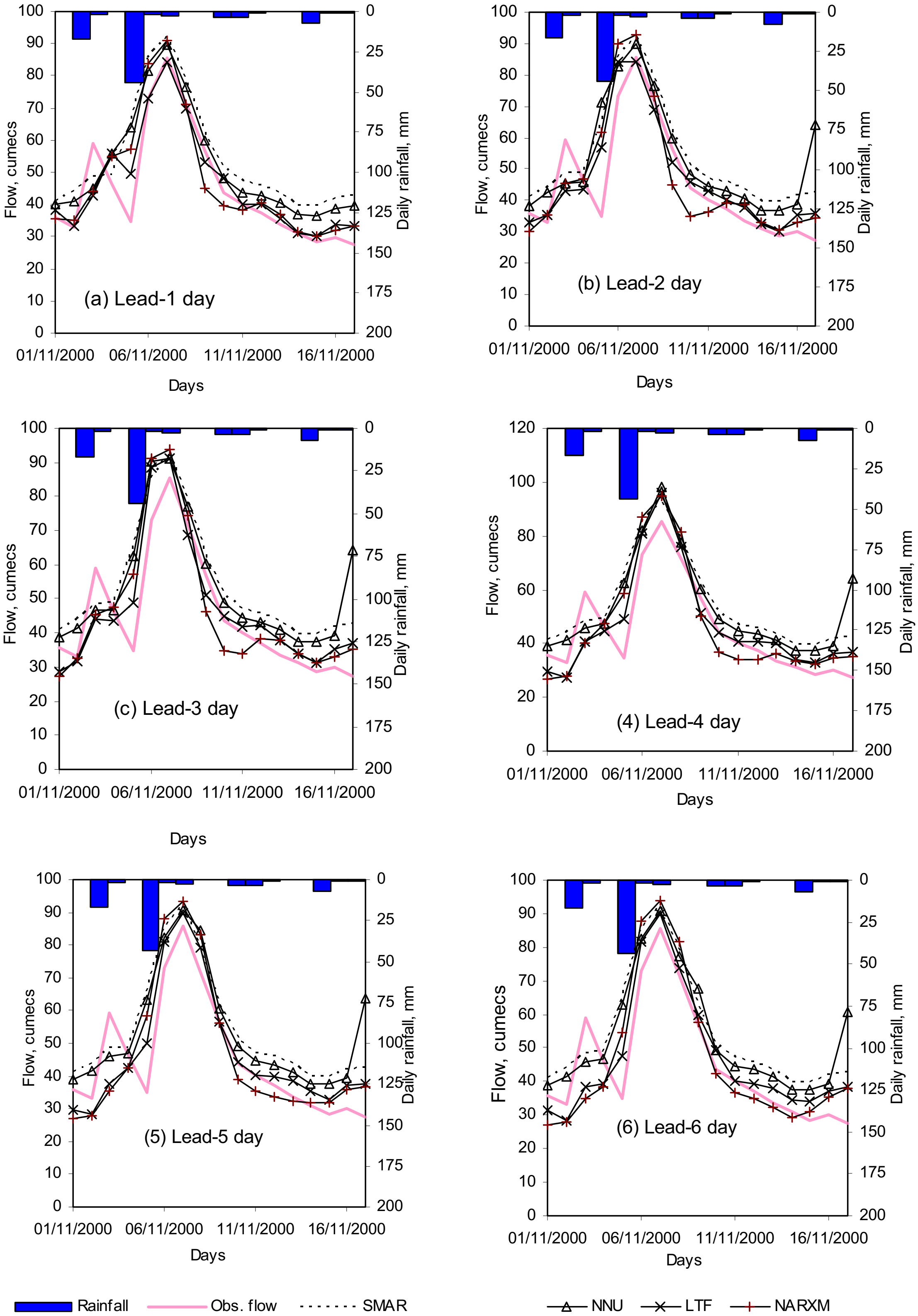

Fig. 4 (a-f). *Comparison of NARXM, LTF and NN-updating model forecasts for (a) lead-1 day to (f) lead-6 days, for the second highest flood, of November 2000 (in the verification period)*

References

Ahsan, Mainul and O'Connor, K.M., 1994. A simple nonlinear rainfall-runoff model based on the concept of a variable gain factor. *J. Hydrol.*, **155**, 151–183.

Beven, K. and Pappenberger, F., 2003. Discussion on 'Towards the hydraulics of the hydroinformatics era', by Abbott, M.B., Babovic, V.M., and Cunge, J.A., (in *J. Hydraul. Res.*, **39**, 2001, 339–349). *J. Hydraul. Res.*, **41**, 331–336) .

Bardsley, W.E., 1983. An alternative distribution for describing the instantaneous unit hydrograph. *J. Hydrol.*, **62**, 375–378.

Box, G.E.P. and Jenkins G.M., 1976. *Time series analysis: forecasting and control*. Holden-Day, San Francisco, CA, USA.

Brath, A., Montanari, A. and Toth, E., 2002. Neural networks and non-parametric methods for improving real-time flood forecasting through conceptual models. *Hydrol. Earth Syst. Sci.*, **6**, 627–640.

Ciarapica, L. and Todini, E., 2002. TOPKAPI: A Model for the representation of the rainfall-runoff process at different scales. *Hydrolog. Process.*, **16**, 207–229.

Georgakakos, K.P., 1986. A generalized Stochastic Hydro-meteorological Model for Flood and Flash-Flood Forecasting; 1. Formulation. *Water Resour. Res.*, **22**, 2083–2095.

Goswami, M, O'Connor, K.M. and Shamseldin, A.Y., 2002a. Structures and Performances of Five Rainfall-Runoff Models for Continuous River-Flow Simulation. In: *Integrated Assessment and Decision Support*, Proc. 1st Biennial Meeting Int. Environmental Modelling and Software Society, 24-27 June, 2002, University of Lugano, Switzerland). **1**, 476–481.

Goswami, M, O'Connor, K.M. and Shamseldin, A.Y., 2002b. Rainfall-runoff modelling of two Irish catchments (one Karstic and one non-Karstic). In: *Celtic Water in European Framework, Proc. Third Inter-Celtic Colloquium Hydrology and Management of Water Resources*, Dept. Engineering Hydrology, NUI, Galway, Ireland. 151–164.

Kachroo, R.K., 1992a. River Flow forecasting, Part 5, Applications of a conceptual model. *J. Hydrol.*, **133**, 141–178.

Kachroo, R.K., 1992b. River flow forecasting, Part 1, A discussion of principles. *J. Hydrol.*, **133**, 1–15.

Kachroo, R.K. and Liang, G.C., 1992. River flow forecasting. Part 2. Algebraic development of linear modelling techniques. *J. Hydrol.*, **133**, 17–40.

Kachroo, R.K. and Natale, L., 1992. Non-linear modelling of the rainfall-runoff relation. *J. Hydrol.*, **135**, 341–369.

Kachroo, R.K., Liang, G.C. and O'Connor, K.M., 1988. Application of the linear perturbation model (LPM) to flood routing on the Mekong River. *Hydrol. Sci. J.*, **33**, 193–214.

Khan, H., 1986. *Conceptual modelling of rainfall-runoff systems*. M. Eng. Thesis, National University of Ireland, Galway.

Koren, V., Reed, S., Smith, M. and Zhang, Z., 2003. Combining physically-based and conceptual approaches in the development and parameterization of a distributed system. In: *Weather Radar Information and Distributed Hydrological Modelling. IAHS publlication no. 282*, 101–108.

Koren, V., Reed, S., Smith, M., Zhang, Z. and Seo. D.J., 2004. Hydrology Laboratory Research Modelling System (HL-RMS) of the US National Weather Service. *J. Hydrol.*, **291**, 297–318.

Liang, G.C., 1992. A note on the revised SMAR model. Workshop Memorandum, Department of Engineering Hydrology, National University of Ireland, Galway (Unpublished).

Liang, G.C. and Nash, J.E., 1988. Linear models for river flow routing on large catchments. *J. Hydrol.*, **103**, 157–188.

Liang, G.C., O'Connor, K.M. and Kachroo, R.K., 1994. A multiple-input single-output variable gain factor model. *J. Hydrol.*, **155**, 185–198.

Liu, Z. and Todini, E., 2002. Towards a comprehensive physically-based rainfall-runoff model. *Hydrol. Earth Syst. Sci.*, **6**, 859–881.

Ljung, L. and Soderstrom, T., 1983. *Theory and Practice of Recursive Identification*, MIT Press: Cambridge, Mass., USA.

Michaud, J. and Sorooshian, S., 1994. Comparison of simple versus complex distributed runoff models on a mid-sized semi-arid watershed. *Water Resour. Res.*, **30**, 593–605.

Nash, J.E. and Barsi, B.I., 1983. A hybrid model for flow forecasting on large catchments. *J. Hydrol.*, **65**, 125–137.

Nash, J.E. and Foley, J.J., 1982. Linear models of rainfall-runoff systems. In: *Rainfall-Runoff Relationship*, V.P. Singh (Ed.), *Proc. Int. Symp. on Rainfall-Runoff modelling*. Mississippi State University, May 1981, USA. Water Resources Publications. 51–66.

Nash, J.E. andSutcliffe, J.V. , 1970. River flow forecasting through conceptual models, Part 1, A discussion of principles. *J. Hydrol.*, **10**, 282–290.

Norton, J.P., 1986. *An Introduction to Identification*. Academic press. London

O'Connor, K.M., 1976. A discrete linear cascade model for hydrology. *J. Hydrol.*, **29**, 203–242.

O'Connor, K.M., Goswami, M., Liang, G.C., Katchroo, R.K. and Shamseldin, Y.A., 2001. The Development of the 'Galway Real-Time River Flow Forecasting System (GFFS): Paper No. 035. In: S*ustainable Use of Land and Water*, *Proc. 19th European ICID Conference*, Brno & Prague, Czech Republic, 4-8 June, 2001. ISBN 80-238-7102-1 (Abstracts and CD of full texts).

O'Connell, P.E., 1991. A historical perspective. In: *Recent advances in the modelling of hydrologic systems*, D.S. Bowles and P.E. O'Connell, P.E. (Eds.), Kluwer, The Netherlands. 3–30.

O'Connell, P.E., Nash, J.E. and. Farrell, J.P, 1970. River Flow forecasting through conceptual models. Part 2. The Brosna catchment at Ferbane. *J. Hydrol.,* **10**, 317–329.

Peetanonchai. B., 1995. *A modification of the AR model for updating of model forecasts.* Unpublished Report for the Sixth Intl. Advanced Course/Workshop on River Flow Forecasting (24 April – 30 June, 1995, Dept. of Eng. Hydrology, National University of Ireland, Galway).

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Se, D.J. and DMIP Participants, 2004. Overall distributed model intercomparison project results. *J. Hydrol.*, in Press; (Corrected Proof available on-line since 1st July 2004).

Seyfried, M.S. and Wilcox, B.P., 1995. Scale and the nature of spatial variability: Field examples having implications for hydraulic modelling. *Water Resour. Res.*, **31**, 173–184.

Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modeling. *J. Hydrol.*, **199**, 272–294.

Shamseldin, A.Y. and O'Connor, K.M., 1999. A real-time combination method for the outputs of different rainfall-runoff models. *Hydrol. Sci. J.*, **44**, 895–912.

Shamseldin, A.Y., and O'Connor, K.M., 2001. A non-linear neural network technique for updating of river flow forecasts. *Hydrol. Earth Syst. Sci.*, **5**, 577–597.

Shamseldin, A.Y., O'Connor, K.M. and Liang, G.C., 1997. Methods for combining the outputs of different rainfall-runoff models. *J. Hydrol.*, **197**, 203–229.

Smith, M., Seo, D.J., Koren, V., Reed, S., Zhang, Z., Moreda, F. and Kuzmin, V., 2003. *Results of the Distributed Model Intercomparison Project (DMIP)*. Hydrological Sciences Programme HS 18, Incorporating hydrological process knowledge into catchment modelling. EGS-AGU-EUG Joint Assembly, Nice, France, 06-11 April 2004. Page 169.

Tan, B.Q. and O'Connor, K.M., 1996. Application of an empirical infiltration equation in the SMAR conceptual model. *J. Hydrol.*, **185**, 275–295.

Wheater, H.S., Jakeman, A.J. and Beven. K.J., 1993. Progress and directions in rainfall-run-off modelling. In: *Modelling Change in Environmental Systems*, A.J. Jakeman, M.B. Beck and M.J. McAleer (Eds.), Wiley, Chichester, UK.

WMO 1992, *Simulated Real-Time Intercomparison of Hydrological Models*. OHR-38. WMO-No. 779. Geneva Switzerland.

Woolhiser, D.A., 1996. Search for a physically based runoff model—a hydrologic El Dorado? *J. Hydraul. Eng. ASCE* **122**, 122-129.

Xiong, L, O'Connor, K.M. and Goswami M., 2001. Application of the Artificial Neural Network (ANN) in Flood Forecasting on a Karstic Catchment. Conference Paper. In: *Forecasting and Mitigation of Water-Related Disasters*, Proc. XXIX Congress of the International Association of Hydraulic Engineering and Research (IAHR), Beijing, China, 16–21 September, 2001.

Ye, W., Bates, B.C., Viney, N.R., Sivapalan, M. and Jakeman, A.J., 1997. Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments. *Water Resour. Res.*, **33**, 153–166.

Young, P.C., 1974. Recursive Approach to Time Series Analysis, *Bull. Inst. Math. Appl.*, **12**, 209–224.

Young, P. C., 1984. *Recursive Estimation and Time Series Analysis: An Introduction*, Springer-Verlag, Berlin, Germany.

Young, P.C., 2002 Advances in real-time flood forecasting. *Phil. Trans. Roy. Soc., Phys. Eng. Sciences,* **360**, 1433–1450.

Young, P. and Jakeman, A. 1979. Refined instrumental variable methods of recursive time-series analysis. Part 1. Single input, single output systems. *Int. J. Control*, **29**, 1–30.

Zhang, J.Y., O'Connor, K.M. and Liang, G.C., 1994. A software package for river flow forecasting based on the SMAR model. In: *Water resources and distribution*, W.R. Blain and K.L. Katsifarakis (Eds.), Hydraulic Engineering Software V Vol.1, Computational Mechanics Publications, Ashurst Lodge, Ashurst, Southampton, UK, 163–170.

APPENDIX: The GFMFS Rainfall-Runoff models (which operate in 'simulation' mode)

SIMPLE LINEAR MODEL (NP-SLM AND P-SLM)

The intrinsic hypothesis of the naïve/primitive Simple Linear Model (SLM) is the assumption of a linear time-invariant relationship between the total rainfall and the total discharge. In its discrete form, the non-parametric simple linear model (NP-SLM) is expressed by the well-known convolution summation relation,

$$Q_t = \sum_{j=1}^{m} R_{t-j+1} h'_j + e_t \;=\; G \sum_{j=1}^{m} R_{t-j+1} b_j + e_t, \quad \text{with } \sum_{j=1}^{m} h'_j = G, \;\; b_j = {h'_j}/{G}, \text{ and } \sum_{j=1}^{m} b_j = 1, \tag{A.1}$$

in which Q_i and R_i are the discharge and rainfall respectively at the t^{th} time-step, h'_j is the j^{th} discrete pulse response ordinate (weight), m is the memory length of the system, e_t is the forecast error term and G is the Gain Factor which reflects the ratio of the total volume of the observed discharge to that of the observed rainfall input (Kachroo and Liang, 1992).

In its parametric form, the Simple Linear Model (P-SLM) has a Linear Transfer Function (LTF)-type representation of the transformation process of the input series R_t to the output series Q_t, for discrete data intervals, also referred to as an Autoregressive Exogenous-input (ARX) type of model, which can be written as the linear difference equation:

$$Q_t = \sum_{j=0}^{r} \alpha_j Q_{t-j} + \sum_{j=1}^{s} \omega_j R_{t-\beta-j+1} + e_t \tag{A.2}$$

in which α_j and ω_j are the autoregressive and exogenous-input parameters respectively, which are conveniently estimated directly by the method of Ordinary Least Squares (OLS), β is the pure time delay restricted to integer values only, r and s are the orders respectively of the autoregressive and exogenous-input parts, and e_t is the model forecast error term. Although the model is calibrated in updating mode, when operating in the *non-updating* (simulation) mode, only previously *computed* values of Q, i.e. $\hat{Q}_{t-1,}\ \hat{Q}_{t-2}$, etc. with the exception of e_t assumed to be zero, are used on the right hand side of the above transfer function Eqn. (A.2).

LINEAR PERTURBATION MODEL (NP-LPM AND P-LPM)

In the linear perturbation model (Nash and Barsi, 1983; Kachroo *et al.*, 1988; Liang and Nash, 1988; Kachroo and Liang, 1992), it is assumed that, during a year in which the daily rainfall is identical to its seasonal daily expectation, the corresponding discharge hydrograph is also identical to its seasonal expectation. However, in all other years, when the rainfall and the discharge values depart from their respective seasonal expectations, these departures (perturbation) series are assumed to be related by a discrete linear time invariant system (see Fig. A.1).

In the application of the LPM, it is necessary to obtain an estimate of the expected values of the input (or rainfall) and discharge for each date, i.e. day of the year 1–365. Smoothing of the season mean series is done by Fourier (i.e. harmonic) analysis (Kachroo *et al.*, 1988)

For the non-parametric linear perturbation model (NP-LPM), the relation between the two perturbation series may be represented algebraically by the convolution summation equation. Hence, this *linear* component can be calibrated by the *OLS* method. Although overall, the NP-LPM is *non-linear*, being *seasonally-based*, it can be regarded as a *'quasi-linear'* model.

The mathematical form of the *linear* component of the Parametric Linear Perturbation Model (P-LPM) is identical in structure to that of the P-SLM, i.e. the LTF structure. The only difference is that the daily series of departures of inputs and outputs from their daily seasonal mean values are used in the transfer function equation for the P-LPM instead of the total rainfall and discharge series as used in the P-SLM.

THE LINEARLY-VARYING GAIN FACTOR MODEL (LVGFM)

The LVGFM (see Fig. A.2), which is also a *'quasi-linear'* model, involves only the variation of the Gain Factor G with a selected index of the prevailing catchment wetness, without varying the shape (i.e. the weights) of the normalised response function (Ahsan and O'Connor, 1994). The model has the familiar convolution summation structure, but is based on the concept of a *time-varying gain factor* G_t, i.e.

$$Q_t = G_t \sum_{j=1}^{m} R_{t-j+1} B_j, \qquad \text{where} \quad \sum_{j=1}^{m} B_j = 1 \qquad \text{(A.3)}$$

In its simplest form, G_t is *linearly* related to an index of the soil moisture state z_t of the catchment by the equation $G_t = G_t(z_t) = a + bz_t$, where a and b are parameter constants, but more complex $G_t(z_t)$ relations have also been used (Ahsan and O'Connor, 1994). The values of z_t are conveniently obtained from the outputs of an auxiliary rainfall-runoff model, such as the naïve SLM, the form of z_t for the case of the SLM being

$$z_t = \frac{\hat{Q}_t}{\overline{Q}} = \frac{\hat{G}}{\overline{Q}} \sum_{j=1}^{m} R_{t-j+1} \hat{h}_j \qquad \text{(A.4)}$$

where $\hat{G}$ and $\hat{h}_j$ are estimates of the SLM Gain Factor and pulse response ordinates respectively and $\overline{Q}$ is the mean discharge in the calibration period. However, the output $\hat{Q}_t$ of any other selected auxiliary model (e.g. SMAR) could likewise be used for this purpose of estimating z_t. The overall operation of the LVGFM, as used in this study, has the mathematical form

$$\begin{aligned} Q_t &= a\sum_{j=1}^{m} R_{t-j+1}B_j + b\, z_t \sum_{j=1}^{m} R_{t-j+1}B_j + e_i = \sum_{j=1}^{m} R_{t-j+1}\left(aB_j\right) + \sum_{j=1}^{m}\left(z_t R_{t-j+1}\right)\left(bB_j\right) + e_t \\ &= \sum_{j=1}^{m} R_{t-j+1}B'_j + \sum_{j=1}^{m} R''_{t-j+1}B''_j + e_t \qquad \text{(A.5)} \end{aligned}$$

where $B'_j = a\,B_j$, $R''_{t-j+1} = z_t R_{t-j+1}$, $B''_j = b\,B_j$, and $\sum B_j = 1.0$.

The *effective* parameter sets B'_j and B''_j can be estimated conveniently by the *OLS* method.

ARTIFICIAL NEURAL NETWORK MODEL (ANNM)

The *ANN* provides a flexible non-linear mapping of the inputs into the outputs without specifying *a priori* the mathematical nature of the relation between inputs and outputs. However, it must be configured in advance of training/calibration, i.e. the structure of the network must be specified. The GFMFS has adopted the 'multi-layer feed-forward' type of artificial neural network which consists of an input layer, an output layer and only one 'hidden' layer located between the input and the output layers, as used by Shamseldin (1997) for rainfall-runoff modelling.

Neural networks are by nature 'parameter rich'; their parameters have 'no clear physical interpretation' and uncertainty inevitably arises in their identification and in the use of the network for prediction (Beven and Pappenberger, 2003). In spite of these drawbacks, the use of neural networks in hydrology has proliferated in recent years, often being used blindly and without question.

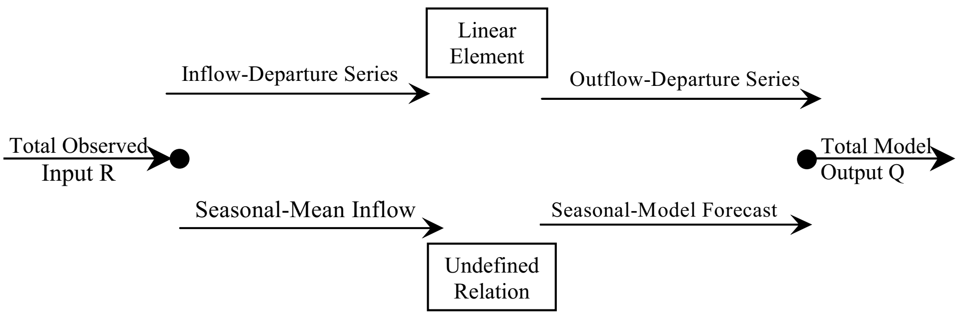

Fig. A.1: Schematic diagram of the Linear Perturbation Model (LPM)

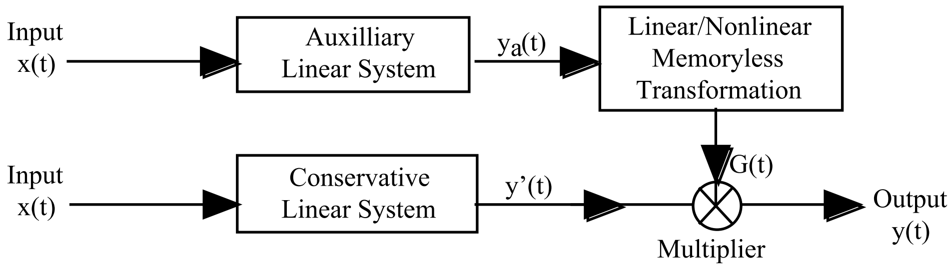

Fig. A.2: Schematic diagram of the Linearly Varying Gain Factor Model (LVGFM)

Neural networks do not account explicitly for storage/ memory effects! Hence, if only the current and preceding exogenous inputs, i.e. rainfalls, with or without corresponding evaporation data, are used as inputs to the network, it would require a number of preceding rainfalls approximately equal to the effective memory length in order to mimic the storage effects of the catchment, thereby producing a non-parsimonious model. Following the approach of Shamseldin (1997), the number of rainfall inputs required can be reduced drastically by providing at least the current forecast of discharge provided by an auxiliary rainfall-runoff model, such as the naïve SLM (or preferably a better model such as SMAR), perhaps with one or more of the preceding discharge forecasts, as inputs to the network, thereby reducing the number of input neurons. (Note that if preceding observed discharges are used as inputs, instead of the forecasted values suggested above, then the network would function as an updating model.)

In the GFMFS, for the results presented in this study, the Simplex search technique was used for training the network.

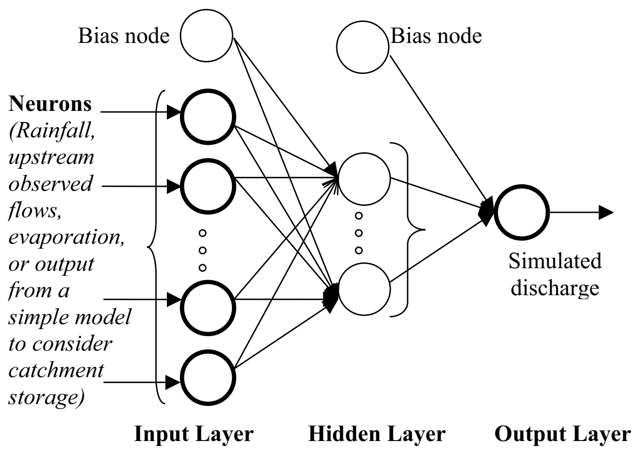

Fig. A.3. *Schematic diagram of the Artificial Neural Network (ANN) model*

SOIL MOISTURE ACCOUNTING AND ROUTING (SMAR) MODEL

The Soil Moisture Accounting and Routing (SMAR) model is a lumped rainfall–evaporation–discharge model of the conceptual type, which has been developed from the 'Layers' conceptual rainfall-runoff model of O'Connell *et al.* (1970). In this model, the input variables, i.e. rainfall and evaporation, are transformed into simulated discharge through a series of steps which, in a very simplified manner, mimic the dominant physical processes (excluding snowmelt) in the rainfall-runoff transformation.

Using a number of empirical and assumed relations which are considered to be at least physically plausible, the non-linear water balance (i.e. soil moisture accounting) component of SMAR model ensures satisfaction of the continuity equation, over each time-step, i.e. it preserves the balance between the rainfall, the evaporation, the 'generated runoff' (which eventually, after routing, contributes to the simulated runoff) and the changes in the various elements (layers) of soil moisture storage. The routing component, on the other hand, simulates the attenuation and the diffusive effects of the catchment by routing the various generated runoff components through linear time-invariant storage elements. For each time-step, the combined output of the two routing elements adopted (i.e. one for the sum of the generated 'surface runoff' components and the other for the generated 'groundwater runoff') becomes the simulated discharge forecast. In the GFMFS package, three two-parameter-distribution options are available for routing the generated 'surface runoff' component, namely the classic Gamma (Nash IUH) (the default method), the Negative Binomial (Pascal) (O'Connor, 1976) and (for sharp-peaked responses) the Inverse Gaussian distribution (Bardsley, 1983). A single discrete linear reservoir is used as the routing component for the generated 'groundwater runoff'.

Two variants of the original SMAR model have been incorporated in the GFMFS, the default 9-parameter

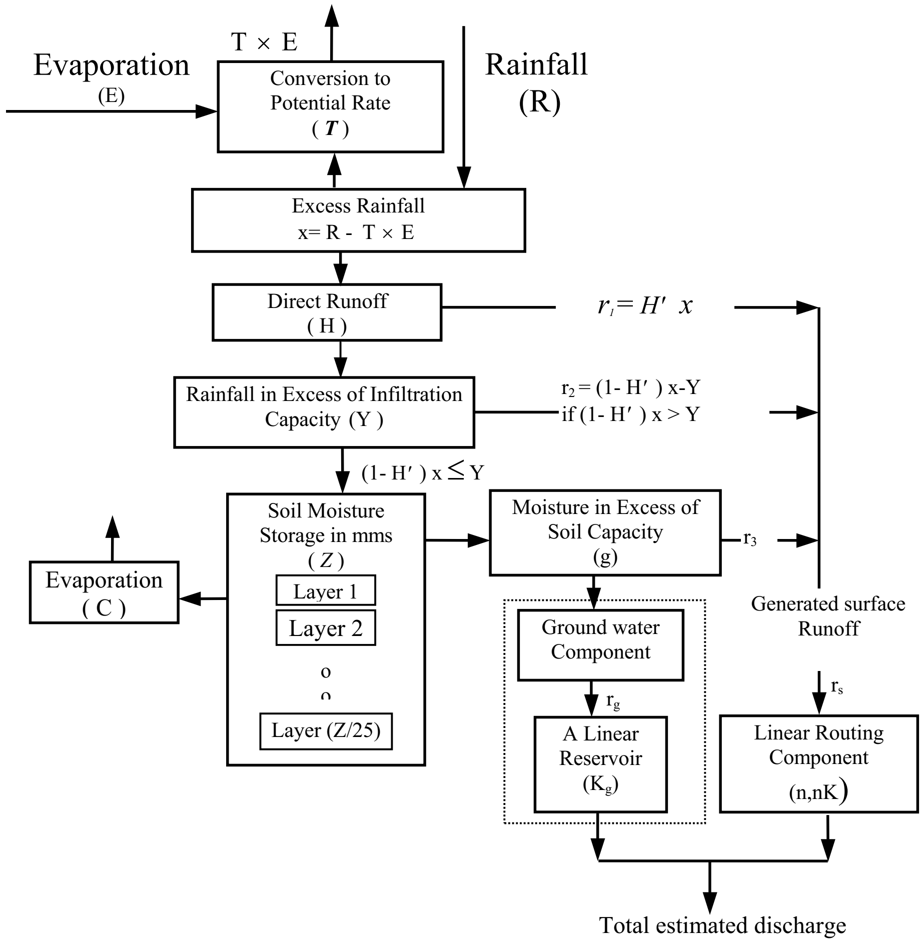

Fig. A.4. *Structure of the 9-paremeter SMAR conceptual model*

SMARG model (refer to Fig. A.4) (with five water balance parameters, *T, H, Y, C,* and *Z*; one weighting parameter for groundwater routing, *G*; and three routing parameters, *N, NK* and K_B), and the 10-parameter SMARK model devised for application on a karstic catchment, the tenth parameter *F* being the coefficient for loss (or gain) from the ground water storage.

The choice of three automatic optimisation algorithms, i.e. the Genetic Algorithm, the Rosenbrock Direct-Search Method and the Simplex Method, are available for the calibration of the SMAR conceptual model. There is also the option of sequential optimisation, starting with the Genetic Algorithm, then the Rosenbrock Method and finally the Simplex Method.